



# UNITED STATES PATENT AND TRADEMARK OFFICE

54  
UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER OF PATENTS AND TRADEMARKS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/602,262	06/23/2000	Paul S. Cohen	YOR9-2000-0131-US1	9324

7590 06/02/2003

Anne Vachon Dougherty  
3173 Cedar Road  
Yorktown Heights, NY 10598

EXAMINER

LERNER, MARTIN

ART UNIT	PAPER NUMBER
----------	--------------

2654

DATE MAILED: 06/02/2003 5-

Please find below and/or attached an Office communication concerning this application or proceeding.

54

# Office Action Summary

Application No.

09/602,262

Applicant(s)

COHEN ET AL.

Examiner

Martin Lerner

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1 to 10 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1 to 10 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on \_\_\_\_ is: a) ☐ approved b) ☐ disapproved by the Examiner.  
If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

## Priority under 35 U.S.C. §§ 119 and 120

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a) ☐ All b) ☐ Some \* c) ☐ None of:  
1. ☐ Certified copies of the priority documents have been received.  
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.  
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).  
\* See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).  
a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

## Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892) 4) ☐ Interview Summary (PTO-413) Paper No(s). \_\_\_\_.
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948) 5) ☐ Notice of Informal Patent Application (PTO-152)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 4. 6) ☐ Other:

## **DETAILED ACTION**

### ***Specification***

1. The disclosure is objected to because of the following informalities:

On page 4, second line from bottom, the period should be deleted after "type)".

On page 5, line 9, Serial No: 09/369,706 should be updated as "now U.S. Patent No. 6,219,640 issued 17 April 2001".

On page 14, second line from bottom, the serial number should be updated as "09/602,452".

Appropriate correction is required.

### ***Claim Objections***

2. Claims 1 to 10 are objected to because of the following informalities:

In claim 1, line 12, "the audio stream" lacks antecedent basis.

In claim 8, lines 13 to 14, "the audio stream" lacks antecedent basis.

Appropriate correction is required.

### ***Claim Rejections - 35 USC § 103***

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the

Art Unit: 2654

invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. Claims 1, 3 to 5, and 7 to 10 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chen* in view of *Braida et al.*

Regarding independent claims 1 and 9, *Chen* discloses a sound-synchronized video method and system, comprising:

“processing a video signal to generate a video output comprising at least one time stamped acoustic identification of the content of the audio associated with the video signal” – codec CD1 separates the digitized video and audio signals into the digital video and speech components; at the video output of codec CD1, a feature extraction module FE1 extracts mouth information visemes contain the mouth shape and mouth location from the decoded video signal; a memory ME1 stores and time stamps mouth information from the feature extraction module FE1 for phoneme-to-viseme identification (column 2, lines 5 to 47; column 4, lines 36 to 41: Figure 1);

“processing an audio signal to generate an audio output comprising at least one [time stamped] acoustic identification of the content of said audio signal” – codec CD1 separates the digitized video and audio signals into the digital video and speech components; a phoneme recognition module PR1 divides the incoming speech components into recognizable phonemes; lookup table LT1 maps phonemes into visemes (column 2, lines 5 to 22; column 4, lines 26 to 35: Figure 1);

“synchronizing the video signal to the audio signal by adjusting at least one of the signals to align at least one acoustic identification from the video signal with a corresponding acoustic identification from the audio stream” – video and audio signal

Art Unit: 2654

that had become unsynchronized are displayed by synchronizing the video frame to produce sound synchronized video (column 4, lines 33 to 63: Figure 2).

Concerning independent claims 1 and 9, *Chen* discloses the video signal is time stamped, but omits time stamping the audio signal. Only one of the audio and video signals is expressly time stamped in *Chen* because visemes are employed as a reference to synchronize the signals. However, it is common in the prior art to assign time stamps to both audio and video data streams for purposes of synchronization to an absolute time reference. *Braida et al.* teaches a related method and system for synchronizing video images to speech elements where time stamps are applied to both audio and video streams. Phone recognition program 44 assigns start and stop times to digital speech samples 32 (column 6, lines 53 to 58), and digital video images also have time stamps which are referenced to the same time (column 12, lines 13 to 29). It would have been obvious to one of ordinary skill in the art to additionally apply time stamps to the audio signals as taught by *Braida et al.* in the synchronization method and system of *Chen* for the purpose of providing an absolute time reference for synchronization.

Regarding claim 3, *Chen* discloses phoneme recognition module PR1 produces visemes ("the audio identification") from the audio signal and feature extraction module FE1 extracts corresponding mouth information visemes from lookup table LT1; the output video is applied to display DI1 together with the audio signal and produces lip synchronization (column 2, lines 11 to 38: Figure 1).

Regarding claims 4 and 10, *Chen* discloses a method and system for processing a video image, comprising:

“extracting at least one image from the video signal” – codec CD1 separates the digitized video and audio signals into the digital video and speech components (column 2, lines 6 to 11);

“detecting at least one feature in said at least one image” – a feature extraction module FE1 extracts mouth information visemes contain the mouth shape and mouth location from the decoded video signal (column 2, lines 21 to 39: Figure 1);

“analyzing the parameters of said feature” – mouth deformation module MD1 receives inputs from the video signal and information from the feature extraction module FE1, and visemes from lookup table LT1 (column 2, lines 21 to 39: Figure 1);

“correlating at least one acoustic identification to the parameters of said feature” – a viseme is selected from lookup table LT1 that matches features extracted by feature extraction module FE1 (column 2, lines 21 to 39: Figure 1).

Regarding claims 5 and 7, *Chen* discloses speech recognition is at the level of phone groups, corresponding to similar mouth shapes (“articulatory type”) rather than individual phonemes (column 3, line 64 to column 4, line 5); similarly, *Braida et al.* processes phones according to context classes (column 8, line 43 to column 9, line 12: Table 2).

Regarding claim 8, *Chen* discloses speech recognition is at the level of phone groups, corresponding to similar mouth shapes (“articulatory type”) rather than individual phonemes (column 3, line 64 to column 4, line 5); similarly, *Braida et al.*

Art Unit: 2654

processes phones according to context classes (column 8, line 43 to column 9, line 12: Table 2); *Chen* discloses feature extraction module FE1 extracts mouth information visemes containing mouth shape ("a facial feature") (column 2, lines 18 to 31).

5. Claims 2 and 6 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chen* in view of *Braida et al.* as applied to claim 1 above, and further in view of *Basu et al.* ('885).

Concerning claim 2, *Braida et al.* discloses a Viterbi search for purposes of phone recognition (column 6, lines 59 to 61; column 7, lines 51 to 53), but omits utilizing a Viterbi search for purposes of synchronization. However, it is well known that a Viterbi algorithm is utilized for both recognition and time warping alignment. *Basu et al.* ('885) teaches a method of aligning phonemes and visemes with a Viterbi algorithm. (Column 1, Lines 53 to 67) It would have been obvious to one having ordinary skill in the art to utilize a Viterbi algorithm as suggested by *Basu et al.* ('885) in the synchronization method and system of *Chen* for the purpose of aligning phonemes and visemes.

Regarding claim 6, *Chen* discloses speech recognition is at the level of phone groups, corresponding to similar mouth shapes ("articulatory type") rather than individual phonemes (column 3, line 64 to column 4, line 5); similarly, *Braida et al.* processes phones according to context classes (column 8, line 43 to column 9, line 12: Table 2).

Art Unit: 2654

**Conclusion**

6. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Wactlar et al. also discloses time stamps for both audio and images (Abstract; column 4, lines 30 to 48).

Basu et al. ('640) discloses related art.


Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (703) 308-9064. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marsha Banks-Harold can be reached on (703) 305-4379. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 872-9314 for regular communications and (703) 872-9314 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-4700.



ml  
May 27, 2003

  
MARSHA D. BANKS-HAROLD  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2600